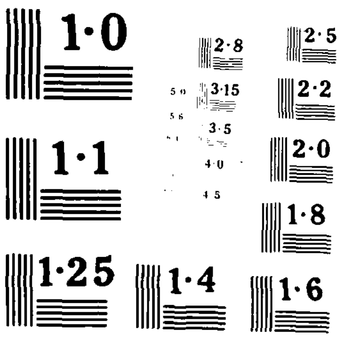
AD-A170 039 A NOTE ON BAYES EMPIRICAL BAYES ESTIMATION BY MEANS OF DIRICHLET PROCESSE (U) STATE UNIV OF NEW YORK AT STONY BROOK DEPT OF APPLIED MATHEMA... L KUO 23 SEP 85 1/1

UNCLASSIFIED AMS-85-60 AFOSR-TR-86-0425 AFOSR-84-0136 F/G 12/1 NL


1·0
2·8
2·5
3·15
2·2
3·5
1·1
2·0
4·0
4·5
1·8
1·25
1·4
1·6

State University of New York at

StonyBrook

Department of Applied Mathematics and Statistics

A Note on Bayes Empirical Bayes Estimation by Means of Dirichlet Processes.

by

Lynn Kuo
SUNY at Stony Brook

Report # AMS 85-60

A Note on Bayes Empirical Bayes Estimation
by Means of Dirichlet Processes.

by

Lynn Kuo
SUNY at Stony Brook

Report # AMS-85-60

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

ADA170039

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S): AMS-85-60	5 MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 86-0425
6a NAME OF PERFORMING ORGANIZATION: S.U.N.Y. at Stony Brook; 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research
6c ADDRESS (City, State and ZIP Code): Dept. of Applied Math. and Stat. S.U.N.Y., Stony Brook, N.Y.11794	7b ADDRESS (City, State and ZIP Code): Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448
8a NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b OFFICE SYMBOL (If applicable): NM	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-84-0136
8c ADDRESS (City, State and ZIP Code): Bolling AFB DC 20332-6448	10 SOURCE OF FUNDING NOS: PROGRAM ELEMENT NO 61102F; PROJECT NO 2304; TASK NO A5; WORK UNIT NO

11 TITLE (Include Security Classification): A NOTE ON BAYES EMPIRICAL BAYES ESTIMATION BY MEANS OF DIRICHLET PROCESSES.

12 PERSONAL AUTHOR(S): Lynn Kuo

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Yr, Mo, Day)	15 PAGE COUNT
Research Report	FROM 7/84 TO 9/85	1985, September, 23	17

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB GR	Dirichlet Processes, Mixtures of Dirichlet Processes, Bayes Empirical Bayes Estimation.

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

Bayes estimators are derived by means of the Dirichlet process hyperprior approach for general empirical Bayes problems. For any sample size, these estimators are expressed concisely as ratios of two multidimensional integrals. A numerical example on Poisson Sampling is given.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS ☐	21 ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL: Brian Woodruff	22b TELEPHONE NUMBER (Include Area Code): (202) 767- 5027; 22c OFFICE SYMBOL: NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

A Note on Bayes Empirical Bayes Estimation by Means of Dirichlet Processes

by

Lynn Kuo
State University of New York,
Stony Brook

Abstract

Bayes estimators are derived by means of the Dirichlet process hyperprior approach for general empirical Bayes problems. For any sample size, these estimators are expressed concisely as ratios of two multidimensional integrals. A numerical example on Poisson sampling is given.

Abbreviated Title

Bayes Empirical Bayes Estimation

AMS 1980 Subject Classifications

Primary 62C12, Secondary 62G05

Key Words and Phrases

Dirichlet process, mixtures of Dirichlet processes, Bayesian non-parametric density method, Bayes empirical Bayes estimation, compound Poisson distribution.

Lynn Kuo, Statistical Survey Institute, Statistical Research Division, Statistics Reporting Service, USDA, Washington, D.C. 20250.

This work is partially supported by a grant from the Office of Scientific Research, United States Air Force 1984-1986.

1. Introduction

In the general setting of empirical Bayes problems, it is assumed that the unobservable parameters $\{\theta_i\}$, $i = 1, \ldots, n$, are taken independently from an unknown distribution G, and that associated with each θ_i, a random variable x_i is observed with known probability density $f(x_i|\theta_i)$ with respect to some σ-finite measure ν on the real line. It is also assumed, given the $\{\theta_i\}$, the observations x_i, $i = 1, \ldots, n$, are independent. We intend to make inferences about $\{\theta_i\}$ or G from the observations.

Several approaches to estimating G or $\{\theta_i\}$ are available. One approach is to use the observed data to estimate the mixing distribution G and use this estimated G as a Bayes prior. To estimate this prior, most authors assume a parametric representation of the prior with unknown parameters estimated by the data. There is another approach to empirical Bayes problems, namely the Dirichlet process hyperprior approach, where the $\{\theta_i\}$ are taken independently from a random distribution G which is chosen from the Ferguson's Dirichlet process (1973) indexed by a finite measure α. The measure α usually represents the statistician's prior belief about G. The Bayes estimator of G or $\{\theta_i\}$ can be derived. As pointed out by Anderson and Louis (1979), this approach is potentially superior, since the construction of estimators does not depend on a specific form for the prior. Another desirable feature is pointed out by Berger (1980b, p. 83): the statistician can combine subjective information α and past data to estimate G and $\{\theta_i\}$, unlike the usual empirical Bayes approach, where the unknown prior parameters are completely estimated by the data.

This Dirichlet hyperprior approach was first proposed by Antoniak (1974) and subsequently studied by Berry and Christensen (1979), Anderson and Louis (1979). The usefulness of this approach had been limited in the past by the following two deficiencies. 1. No concise expressions for the proper Bayes estimators had been given for $n > 3$ due to the complex bookkeeping and labor involved in deriving them. 2. No satisfactory numerical methods had been developed in evaluating those estimators. With the work of Lo (1978), Bayes estimators can be derived for any n. The purpose of this note is to exhibit Bayes estimators of $\{\theta_i\}$ for arbitrary n. It can be seen from equation (1) of Section 2 that each of the Bayes estimators can be written as a ratio of n-dimensional integrals. These integrals are hard to evaluate explicitly due to the high dimensionality and the fact that the integrands are peaked in a small region of the parameter space. In a recent article, Kuo (1985) proposes to circumvent this problem by 1. decomposing each of the multidimensional integrals into a weighted average of products of one-dimensional integrals and 2. approximating each of the weighted averages by an importance sampling Monte Carlo method. It is easy to implement this method. Moreover, the Monte Carlo estimator has been demonstrated to work well in terms of efficiency and precision. For the detailed method, statistical analysis and numerical examples, see Kuo (1985). A numerical example on Poisson sampling using the method of Kuo is given here.

See Robbins (1955) for the pioneer development of empirical Bayes methods of estimation. See Susarla (1982) for an expository article on empirical Bayes theory which also includes some of the recent developments in this area.

2. Derivation of Bayes Estimators

To derive the Bayes rule of $\{\theta_i\}$, let us make the following assumptions:

i. Let α be a finite measure with finite second moment on a measurable space $(\mathbb{R}, \mathcal{B})$ with $\mathbb{R}$ the real line and $\mathcal{B}$ the σ-field of Borel sets. An unknown distribution G is chosen from a Dirichlet process with parameter α.

ii. Given G, the unobservable $\theta_1, \ldots, \theta_n$ are chosen independently from G.

iii. Given G and $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_n)$, the observations $\underset{\sim}{x} = (x_1, \ldots, x_n)$ have density $f(\underset{\sim}{x} \mid \underset{\sim}{\theta}) = \prod_{i=1}^{n} f_i(x_i \mid \theta_i)$ idependent of G, where for all $i = 1, \ldots, n$, f_i is a density dominated by ν and $f_i(x \mid \theta)$ is measurable in θ for all x.

iv. The loss function is given by

$$L(\underset{\sim}{\theta}, \underset{\sim}{a}) = \sum_{i=1}^{n} (\theta_i - a_i)^2 .$$

<u>REMARK 1.</u> Note that in Assumption iii, we allow different f_i in the model. This has the advantage of incorporating individual characteristics of the x_i's in the model, such as combining normal and gamma components (see Berger, 1980a), or treating x_i with different variances, etc.

<u>REMARK 2.</u> Note that given G, the x_i are independently distributed according to $\int f_i(x_i|\theta)G(d\theta)$. This is essentially the random density considered by Lo. The main difference is that his objective is to estimate the mixing distribution G and various functions of G; our objective is to estimate the θ_i's.

REMARK 3. In addition to Lo's work, nonparametric density estimation has also been studied by Ferguson (1983). The Monte Carlo method described in Section 1 was adapted by Ferguson to compute the density estimator. The feasibility of this method and error reduction techniques were further illustrated.

REMARK 4. Empirical Bayes estimation (as opposed to Bayes estimation) of the density function $f(x|G) = \int f(x|\theta)G(d\theta)$ described in Remark 2 has also been studied by Ghorai and Susarla (1982).

Before proving the main result, we first define some notation and state a lemma. Let α and G be defined as in Assumption i. Then P_α denotes the probability measure on $(\mathfrak{B}, A)$ yielding the random distribution G, where $\mathfrak{B}$ is the space of distribution functions on $(\mathbb{R}, B)$, and A is the σ-field of Borel sets in the Levy metric. The following lemma is from Lo (1978, 1984).

LEMMA 1. Let θ_1 be chosen from G. If $g(\theta_1, G)$ is a quasi-integrable function with respect to the joint probability $G(d\theta_1)P_\alpha(dG)$ defined on $(\mathbb{R} \times \mathfrak{B}, B \times A)$, then

$$\int_{\mathfrak{B}} \int_{\mathbb{R}} g(\theta_1, G)G(d\theta_1)P_\alpha(dG) = \int_R \int_{\mathfrak{B}} g(\theta_1, G)P_{\alpha+\delta_{\theta_1}}(dG)\frac{\alpha(d\theta_1)}{\alpha(\mathbb{R})} .$$

We are now ready to exhibit the Bayes estimator $\hat{\underset{\sim}{\theta}}$.

THEOREM 1. Given the Assumptions i through iv, the Bayes estimator is given by $\hat{\underset{\sim}{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_n)$, where

$$\hat{\theta}_k = E(\theta_k | x_1, \ldots, x_n) = \frac{\int_{\mathbb{R}_n} \cdots \int \left(\theta_k \prod_{i=1}^{n} f_i(x_i|\theta_i)\right) \prod_{i=1}^{n} \left\{\alpha + \sum_{j=1}^{i-1} \delta_{\theta_j}\right\}(d\theta_i)}{\int_{\mathbb{R}_n} \cdots \int \left(\prod_{i=1}^{n} f_i(x_i|\theta_i)\right) \prod_{i=1}^{n} \left\{\alpha + \sum_{j=1}^{i-1} \delta_{\theta_j}\right\}(d\theta_i)} \tag{1}$$

for all $k = 1, \ldots, n$.

Proof: Let $h(\underset{\sim}{x},\underset{\sim}{\theta})$ denote the joint density of $\underset{\sim}{x}$ and $\underset{\sim}{\theta}$, $\tilde{h}(\underset{\sim}{x})$ denote the marginal density of $\underset{\sim}{x}$, and $\rho_G(\underset{\sim}{\theta})$ denote the joint distribution of θ's given G. Let P_α, $\mathbb{H}$ be defined as in Lemma 1. Then the Bayes estimator under squared error loss for θ_k is given by $\hat{\theta}_k = E(\theta_k \mid \underset{\sim}{x})$, where

$$E(\theta_k|\underset{\sim}{x}) = \frac{\int \theta_k \, h(\underset{\sim}{x},\underset{\sim}{\theta})d\underset{\sim}{\theta}}{\tilde{h}(\underset{\sim}{x})}$$

$$= \frac{\int_{\mathbb{H}} \int_{\mathbb{R}_n} \theta_k \, f(\underset{\sim}{x}|\underset{\sim}{\theta})\rho_G(d\underset{\sim}{\theta})P_\alpha(dG)}{\int_{\mathbb{H}} \int_{\mathbb{R}_n} f(\underset{\sim}{x}|\underset{\sim}{\theta})\rho_G(d\underset{\sim}{\theta})P_\alpha(dG)}$$

$$= \frac{\int_{\mathbb{H}} \left[\prod_{i\neq k}^{n} \int_{\mathbb{R}} f_i(x_i|\theta_i)G(d\theta_i)\right] \cdot \int_{\mathbb{R}} \theta_k f_k(x_k|\theta_k)G(d\theta_k)P_\alpha(dG)}{\int_{\mathbb{H}} \prod_{i=1}^{n} \int_{\mathbb{R}} f_i(x_i|\theta_i)G(d\theta_i)P_\alpha(dG)}$$

$$= \frac{\int_{\mathbb{R}_n} \cdots \int \left(\theta_k \prod_{i=1}^{n} f_i(x_i|\theta_i)\right) \prod_{i=1}^{n} \frac{\left\{\alpha + \sum_{j=1}^{i-1} \delta_{\theta_j}\right\}(d\theta_i)}{\alpha(\mathbb{R}) + i - 1}}{\int_{\mathbb{R}_n} \cdots \int \left(\prod_{i=1}^{n} f_i(x_i|\theta_i)\right) \prod_{i=1}^{n} \frac{\left\{\alpha + \sum_{j=1}^{i-1} \delta_{\theta_j}\right\}(d\theta_i)}{\alpha(\mathbb{R}) + i - 1}}$$

by repeated use of Lemma 1.

REMARK 5. It was shown by Antoniak (1974), the posterior distribution of G given $\underset{\sim}{x}$ is a mixture of Dirichlet processes:

$$G \mid \underset{\sim}{x} \sim \int \mathcal{D}\left(\alpha + \sum_{i=1}^{n} \delta_{\theta_i}\right) dF_{\underset{\sim}{\theta} \mid \underset{\sim}{x}}$$

where $F_{\underset{\sim}{\theta} \mid \underset{\sim}{x}}$ is the posterior distribution of $\underset{\sim}{\theta}$ given $\underset{\sim}{x}$. It can be seen from Theorem 1 (or Remark 2 of this paper and Theorem 1 of Lo (1978)) that the posterior mixing distribution of $\underset{\sim}{\theta}$ may be written as

$$F_{\underset{\sim}{\theta} \mid \underset{\sim}{x}}(C) = \frac{\int_C \cdots \int \prod_{i=1}^{n} f_i(x_i \mid \theta_i) \prod_{i=1}^{n} \left(\alpha + \sum_{j=1}^{i-1} \delta_{\theta_j}\right)(d\theta_i)}{\int_{R^n} \cdots \int \prod_{i=1}^{n} f_i(x_i \mid \theta_i) \prod_{i=1}^{n} \left(\alpha + \sum_{j=1}^{i-1} \delta_{\theta_j}\right)(d\theta_i)}$$

for all $C \in \mathcal{B}^n$, where $(\mathbb{R}^n, \mathcal{B}^n)$ is the n-fold product measure space of $(\mathbb{R}, \mathcal{B})$.

3. Numerical Example

To illustrate the use of the nonparametric Bayes hyperprior approach to empirical Bayes problems and the way the estimates are influenced by the prior choice, an example is given here.

The data set of this example is taken from Bayesian Reliability Analysis (p. 626) by Martz and Waller. Suppose a 10^5 hour life test has been conducted for each of the eleven production lots of a high reliability device. The numbers of failures were observed to be 0, 1, 0, 0, 1, 2, 0, 1, 0, 0, and 0 respectively. It is assumed

that the event of failures can be modeled by Poisson point processes with intensity rate λ_i for the ith lot. The objective is to estimate λ_i for each lot. Let $\theta_i = \lambda_i \cdot 10^5$. A Dirichlet hyperprior approach to this problem is to assume θ_i are distributed according to an unknown distribution G. G is chosen from the Dirichlet process with prior $\alpha\{(-\infty,t]\} = MG_0(t)$, where G_0 represents the statistician's prior guess of G, and M represents the statistician's strengths of prior belief in G_0. Then the Bayes estimator of θ_k can be obtained from (1), where

$$f_i(x_i|\theta_i) = \frac{e^{-\theta_i}\,\theta_i^{x_i}}{x_i!} ,$$

x_i denotes the number of failures in the 10^5 hour test in the ith lot.

The evaluation of the Bayes rule (1) can be approximated by an importance sampling Monte Carlo method proposed by Kuo (1985). If we choose G_0 a gamma distribution $G(\alpha,\beta)$, i.e.,

$$G_0(d\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\,\theta^{\alpha-1}\,e^{-\beta\theta} \cdot I(\theta > 0) ,$$

then the single integrals (see equations (8) and (9) of Kuo, 1985) contained in each Monte Carlo iteration can be evaluated by using the following identities:

$$\int \theta \prod_{i\epsilon k} f(x_i|\theta)G_0(d\theta) = \frac{\beta^\alpha\,\Gamma(\alpha + 1 + \sum_{i\epsilon k} x_i)}{\Gamma(\alpha)\cdot(|k| + \beta)^{(\alpha + 1 + \sum_{i\epsilon k} x_i)} \cdot \prod_{i\epsilon k} x_i} ,$$

and

$$\int \prod_{i\epsilon k} f(x_i|\theta)G_0(d\theta) = \frac{\beta^\alpha\,\Gamma(\alpha + \sum_{i\epsilon k} x_i)}{\Gamma(\alpha)\cdot(|k| + \beta)^{(\alpha + \sum_{i\epsilon k} x_i)} \cdot \prod_{i\epsilon k} x_i} ,$$

where k is a subset of the index set $\{1, \ldots, 11\}$, $|k|$ denotes the number of indices in k.

If we choose G_0 a uniform distribution on $(0, \theta_0)$, then the single integrals are evaluated by

$$\int \theta \prod_{i \in k} f(x_i|\theta) G_0(d\theta) = \frac{I(k\theta_0, 2 + \sum_{i \in k} x_i)\, \Gamma(2 + \sum_{i \in k} x_i)}{\theta_0 \cdot |k|^{(2 + \sum_{i \in k} x_i)} \cdot \prod_{i \in k} x_i!} , \qquad (2)$$

and

$$\int \prod_{i \in k} f(x_i|\theta) G_0(d\theta) = \frac{I(k\theta_0, 1 + \sum_{i \in k} x_i)\, \Gamma(1 + \sum_{i \in k} x_i)}{\theta_0 \cdot |k|^{(1 + \sum_{i \in k} x_i)} \cdot \prod_{i \in k} x_i\, !} ,$$

where $I(y, r) = \int_0^y t^{r-1} e^{-t}\, dt/\Gamma(r)$ denotes the incomplete gamma function.

In the following tables, G_0 is chosen to be either a gamma $G(\alpha,\beta)$ or a uniform distribution $U(0,\theta_0)$. The Bayes rules (1) for various values of α, β, θ_0 and M are computed. They are evaluated by Monte Carlo methods with the number of iterations $NI = 4000$ or $NI = 16000$. Therefore, the posterior standard errors (see equation (12) of Kuo, 1985) are included in parantheses. We have also rearranged the order of the observations for easier visual examination. The intensity rate λ_i are estimated by $\hat{\lambda}_i = \hat{\theta}_i/10^5$.

In Table 1, α, β for the gamma distribution prior guess are chosen to be $\hat{\alpha} = \max\{\bar{x}^2/(S_x^2 - \bar{x}), 0\}$, and $\hat{\beta} = \max\{\bar{x}/(S_x^2 - \bar{x}), 0\}$, where $\bar{x} = \sum x_i/11$, and $S_x^2 = \sum(x_i - \bar{x})^2/10$. This choice of prior

is motivated by assuming that the θ_i's are independent and identically distributed according to $G(\alpha,\beta)$, where α and β are unknown. Considering the marginal distribution of x, we have $E\,x_i = \alpha/\beta$, and $V(x_i) = \alpha(1 + 1/\beta)/\beta$ for all i. The estimators $\hat{\alpha}$ and $\hat{\beta}$ above are obtained by the method of moments and adjusted for $\alpha > 0$ and $\beta > 0$.

Robbins (1980) discussed two methods of obtaining empirical Bayes estimators and proposed a method for combining the two to achieve both consistency and efficiency. The first estimator for θ_i denoted by δ_i, is derived from a nonparametric prior point of view. Then $\delta_i(\underset{\sim}{x}) = (x_i + 1) \cdot \#(x_i + 1)/\#(x_i)$, where $\#(x)$ denotes the number of lots with x items failed. The second estimator for θ_i, denoted by τ_i, is derived from a parametric gamma prior:

$$\tau_i(\underset{\sim}{x}) = (x_i + \hat{\alpha})/(1 + \hat{\beta}) = x_i - (\bar{x}/S_x^2) \cdot I(\bar{x} < S_x^2)\,(x_i - \bar{x}) ,$$

where $x_i \geqslant 0$ for all i. These are Stein type shrinkage estimators. Both types of estimators are given in Table 1. For τ_i, we use the unbiased estimator for the variance of the x's instead of the usual sample variance estimator suggested by Robbins. Otherwise, the shrinkages of the data will pass the origin without the positive part corrections given in the τ_i expression above.

For the Dirichlet hyperprior approach, with the G_0 chosen as $G(\hat{\alpha},\hat{\beta})$ above, we would expect the proper Bayes estimators to approach to the τ's as $M \longrightarrow \infty$. This is confirmed by the estimators shown in Table 1.

In Tables 2(a) and 2(b), a different shape of the gamma distribution is chosen. The mean of G_0 in Table 2(a) is the same as before, i.e., $\bar{x}$. However, the variance of G_0 is 4.54544 which is much larger than 0.0181817, the one used in Table 1. It is interesting to observe that the estimators are closer to the maximum likelihood estimator as $M \longrightarrow \infty$. This reveals that choosing the prior on a mixing distribution is quite different from the usual nonparametric prior situation without the mixing situation. As $M \longrightarrow \infty$, we expect the θ's are more distinct. When $M=500$, $P(\text{all } \theta\text{'s are distinct}) = \prod_{i=1}^{11} (1 - (i-1)/(M+i-1)) = 0.897$. Therefore, the estimator θ_i can be approximated by using x_i alone and the prior. When M is small, many of the θ's are identical with high probability. This explains the phenomenon exhibited in Table 2(a) and 2(b).

The G_0's in Tables 3(a) and 3(b) have the same mean and different variances. Different means and different variances are selected for Tables 4(a) and 4(b) to examine how the estimates are influenced by the prior guess.

In Tables 5(a) and 5(b), the G_0's are chosen to be the uniform distributions $U(0,0.3)$ and $U(0,1)$ respectively. The incomplete gamma distributions described in (2) are evaluated by the MDGAM subroutine of the International Mathematical and Scientific Library (IMSL).

TABLE 1

Bayes and Empirical Bayes Estimates

for $\{\theta_i\}$ with $G_0 \sim G(\hat{\alpha}, \hat{\beta})$

$\hat{\alpha} = 11.3636$ $\hat{\beta} = 25$ NI = 4000

x_i	$\hat{\theta}_i$ (M=1)	$\hat{\theta}_i$ (M=50)	$\hat{\theta}_i$ (M=100)	$\hat{\theta}_i$ (M=500)	ρ_i	τ_i
0	0.4441 (0.0005)	0.4372 (0.0002)	0.4372 (0.0001)	0.4370 (0.0001)	0.4286	0.4371
0	0.4453 (0.0005)	0.4372 (0.0002)	0.4371 (0.0001)	0.4370 (0.0001)	0.4286	0.4371
0	0.4449 (0.0005)	0.4371 (0.0002)	0.4371 (0.0001)	0.4370 (0.0001)	0.4286	0.4371
0	0.4440 (0.0005)	0.4373 (0.0002)	0.4370 (0.0001)	0.4369 (0.0001)	0.4286	0.4371
0	0.4453 (0.0005)	0.4372 (0.0002)	0.4371 (0.0001)	0.4370 (0.0001)	0.4286	0.4371
0	0.4451 (0.0005)	0.4368 (0.0002)	0.4369 (0.0001)	0.4369 (0.0001)	0.4286	0.4371
0	0.4455 (0.0005)	0.4375 (0.0002)	0.4372 (0.0001)	0.4369 (0.0001)	0.4286	0.4371
1	0.4635 (0.0005)	0.4748 (0.0002)	0.4749 (0.0001)	0.4754 (0.0001)	0.6667	0.4755
1	0.4644 (0.0005)	0.4749 (0.0002)	0.4753 (0.0001)	0.4753 (0.0001)	0.6667	0.4755
1	0.4649 (0.0005)	0.4752 (0.0002)	0.4751 (0.0001)	0.4754 (0.0001)	0.6667	0.4755
2	0.4336 (0.0004)	0.5125 (0.0001)	0.5132 (0.0001)	0.5138 (0.0000)	0.0000	0.5140

TABLE 2

Bayes Estimates for $\{\theta_i\}$

with $G_0 \sim G(\alpha, \beta)$

(a) $\alpha = 0.0454544$ $\beta = 0.1$ NI = 16000

x_i	$\hat{\theta}$ (M=1)	$\hat{\theta}$ (M=50)	$\hat{\theta}$ (M=100)	$\hat{\theta}$ (M=500)
0	0.2930 (0.0101)	0.0590 (0.0017)	0.0512 (0.0012)	0.0430 (0.0002)
0	0.2952 (0.0097)	0.0597 (0.0019)	0.0510 (0.0007)	0.0430 (0.0002)
0	0.2926 (0.0102)	0.0614 (0.0021)	0.0499 (0.0006)	0.0433 (0.0002)
0	0.2935 (0.0099)	0.0607 (0.0019)	0.0493 (0.0006)	0.0433 (0.0002)
0	0.2939 (0.0101)	0.0605 (0.0017)	0.0513 (0.0012)	0.0430 (0.0002)
0	0.2901 (0.0096)	0.0576 (0.0015)	0.0511 (0.0012)	0.0430 (0.0002)
0	0.2955 (0.0102)	0.0563 (0.0014)	0.0514 (0.0007)	0.0429 (0.0002)
1	0.7376 (0.0128)	0.9771 (0.0067)	0.9624 (0.0024)	0.9573 (0.0019)
1	0.7354 (0.0129)	0.9715 (0.0069)	0.9647 (0.0030)	0.9534 (0.0011)
1	0.7307 (0.0130)	0.9751 (0.0067)	0.9617 (0.0023)	0.9545 (0.0012)
2	0.3194 (0.0110)	1.6952 (0.0111)	1.7753 (0.0040)	1.8329 (0.0020)

(b) $\alpha = 0.1$ $\beta = 0.01$ NI = 16000

x_i	$\hat{\theta}$ (M=1)	$\hat{\theta}$ (M=50)	$\hat{\theta}$ (M=100)	$\hat{\theta}$ (M=500)
0	0.3161 (0.0081)	0.1156 (0.0012)	0.1076 (0.0008)	0.1004 (0.0002)
0	0.3227 (0.0081)	0.1163 (0.0013)	0.1084 (0.0007)	0.1004 (0.0003)
0	0.3248 (0.0082)	0.1175 (0.0014)	0.1070 (0.0006)	0.1006 (0.0003)
0	0.3183 (0.0081)	0.1170 (0.0012)	0.1070 (0.0006)	0.1007 (0.0003)
0	0.3258 (0.0084)	0.1174 (0.0012)	0.1080 (0.0008)	0.1003 (0.0003)
0	0.3194 (0.0079)	0.1143 (0.0011)	0.1077 (0.0008)	0.1004 (0.0003)
0	0.3268 (0.0081)	0.1138 (0.0011)	0.1083 (0.0007)	0.1002 (0.0002)
1	0.7163 (0.0115)	1.0536 (0.0026)	1.0733 (0.0016)	1.0860 (0.0009)
1	0.7128 (0.0116)	1.0560 (0.0027)	1.0748 (0.0017)	1.0344 (0.0007)
1	0.7073 (0.0116)	1.0535 (0.0026)	1.0715 (0.0017)	1.0650 (0.0008)
2	0.8632 (0.0094)	1.9248 (0.0051)	1.9947 (0.0029)	2.0552 (0.0014)

TABLE 3

Bayes Estimates for $\{\theta\}$

with $G_0 \sim G(\alpha, \beta)$

(a) $\alpha = 22.7272$ $\beta = 50$ NI = 4000

x_i	$\hat{\theta}_i$ (M=1)	$\hat{\theta}_i$ (M=50)	$\hat{\theta}_i$ (M=100)
0	0.4498 (0.0003)	0.4456 (0.0001)	0.4456 (0.0001)
0	0.4500 (0.0003)	0.4457 (0.0001)	0.4456 (0.0001)
0	0.4501 (0.0003)	0.4457 (0.0001)	0.4456 (0.0001)
0	0.4497 (0.0003)	0.4455 (0.0001)	0.4455 (0.0001)
0	0.4499 (0.0003)	0.4455 (0.0001)	0.4455 (0.0001)
0	0.4498 (0.0003)	0.4456 (0.0001)	0.4455 (0.0001)
0	0.4500 (0.0003)	0.4456 (0.0001)	0.4456 (0.0001)
1	0.4598 (0.0003)	0.4649 (0.0001)	0.4649 (0.0001)
1	0.4598 (0.0003)	0.4647 (0.0001)	0.4651 (0.0001)
1	0.4601 (0.0003)	0.4649 (0.0001)	0.4650 (0.0001)
2	0.4693 (0.0002)	0.4841 (0.0001)	0.4844 (0.0001)

(b) $\alpha = 5.6818$ $\beta = 12.5$ NI = 4000

x_i	$\hat{\theta}_i$ (M=1)	$\hat{\theta}_i$ (M=50)	$\hat{\theta}_i$ (M=100)
0	0.4364 (0.0009)	0.4213 (0.0003)	0.4213 (0.0002)
0	0.4372 (0.0009)	0.4216 (0.0003)	0.4212 (0.0002)
0	0.4375 (0.0009)	0.4216 (0.0003)	0.4211 (0.0002)
0	0.4362 (0.0009)	0.4210 (0.0003)	0.4210 (0.0002)
0	0.4368 (0.0009)	0.4210 (0.0003)	0.4210 (0.0002)
0	0.4358 (0.0009)	0.4211 (0.0003)	0.4208 (0.0002)
0	0.4373 (0.0009)	0.4211 (0.0003)	0.4213 (0.0002)
1	0.4746 (0.0010)	0.4941 (0.0003)	0.4940 (0.0002)
1	0.4744 (0.0009)	0.4935 (0.0003)	0.4947 (0.0003)
1	0.4757 (0.0010)	0.4941 (0.0003)	0.4942 (0.0002)
2	0.5091 (0.0009)	0.5662 (0.0002)	0.5674 (0.0002)

TABLE 4

Bayes Estimates for $\{\theta_i\}$

with $G_0 \sim G(\alpha, \beta)$

(a) $\alpha = 11.3636$ $\beta = 12.5$ NI = 4000

x_i	$\hat{\theta}_i$ (M=1)	$\hat{\theta}_i$ (M=50)	$\hat{\theta}_i$ (M=100)
0	0.7493 (0.0012)	0.8366 (0.0004)	0.8394 (0.0003)
0	0.7502 (0.0012)	0.8371 (0.0004)	0.8392 (0.0003)
0	0.7500 (0.0012)	0.8366 (0.0004)	0.8391 (0.0003)
0	0.7486 (0.0012)	0.8358 (0.0004)	0.8389 (0.0003)
0	0.7496 (0.0012)	0.8361 (0.0004)	0.8391 (0.0003)
0	0.7495 (0.0017)	0.8364 (0.0004)	0.8389 (0.0003)
0	0.7501 (0.0012)	0.8362 (0.0004)	0.8391 (0.0003)
1	0.7851 (0.0015)	0.9093 (0.0004)	0.9120 (0.0003)
1	0.7846 (0.0015)	0.9085 (0.0004)	0.9128 (0.0003)
1	0.7871 (0.0015)	0.9091 (0.0004)	0.9122 (0.0003)
2	0.8232 (0.0018)	0.9818 (0.0004)	0.9857 (0.0003)

(b) $\alpha = 10$ $\beta = 50$ NI = 4000

x_i	$\hat{\theta}_i$ (M=1)	$\hat{\theta}_i$ (M=50)	$\hat{\theta}_i$ (M=100)
0	0.2224 (0.0004)	0.1972 (0.0001)	0.1968 (0.0001)
0	0.2227 (0.0004)	0.1973 (0.0001)	0.1968 (0.0001)
0	0.2229 (0.0004)	0.1974 (0.0001)	0.1968 (0.0001)
0	0.2225 (0.0004)	0.1972 (0.0001)	0.1967 (0.0001)
0	0.2226 (0.0004)	0.1972 (0.0001)	0.1967 (0.0001)
0	0.2221 (0.0004)	0.1972 (0.0001)	0.1966 (0.0001)
0	0.2227 (0.0004)	0.1972 (0.0001)	0.1968 (0.0001)
1	0.2331 (0.0003)	0.2166 (0.0001)	0.2161 (0.0001)
1	0.2333 (0.0003)	0.2165 (0.0001)	0.2163 (0.0001)
1	0.2333 (0.0003)	0.2167 (0.0001)	0.2162 (0.0001)
2	0.2422 (0.0002)	0.2358 (0.0001)	0.2356 (0.0001)

TABLE 5

Bayes Estimates for $\{\theta_i\}$

with $G_0 \sim U(0, \theta_0)$

(a) $\theta_0 = 0.3$ NI = 4000

x_i	$\hat{\theta}_i$ (M=1)	$\hat{\theta}_i$ (M=50)	$\hat{\theta}_i$ (M=100)
0	0.2003 (0.0008)	0.1455 (0.0002)	0.1442 (0.0002)
0	0.2007 (0.0008)	0.1457 (0.0002)	0.1442 (0.0002)
0	0.2013 (0.0007)	0.1458 (0.0002)	0.1442 (0.0002)
0	0.2007 (0.0007)	0.1456 (0.0002)	0.1441 (0.0002)
0	0.2005 (0.0008)	0.1455 (0.0002)	0.1441 (0.0002)
0	0.1997 (0.0008)	0.1454 (0.0002)	0.1439 (0.0002)
0	0.2008 (0.0007)	0.1455 (0.0002)	0.1443 (0.0002)
1	0.2222 (0.0004)	0.1965 (0.0002)	0.1956 (0.0001)
1	0.2225 (0.0005)	0.1963 (0.0002)	0.1958 (0.0001)
1	0.2224 (0.0004)	0.1966 (0.0002)	0.1957 (0.0001)
2	0.2323 (0.0002)	0.2223 (0.0001)	0.2219 (0.0001)

(b) $\theta_0 = 1.0$ NI = 16000

x_i	$\hat{\theta}_i$ (M=1)	$\hat{\theta}_i$ (M=50)	$\hat{\theta}_i$ (M=100)
0	0.4601 (0.0015)	0.4192 (0.0004)	0.4185 (0.0003)
0	0.4627 (0.0014)	0.4195 (0.0004)	0.4188 (0.0003)
0	0.4635 (0.0015)	0.4194 (0.0004)	0.4183 (0.0003)
0	0.4615 (0.0015)	0.4195 (0.0004)	0.4185 (0.0003)
0	0.4639 (0.0014)	0.4194 (0.0004)	0.4182 (0.0003)
0	0.4625 (0.0014)	0.4189 (0.0004)	0.4183 (0.0003)
0	0.4635 (0.0014)	0.4185 (0.0004)	0.4187 (0.0003)
1	0.5679 (0.0011)	0.6059 (0.0003)	0.6067 (0.0002)
1	0.5673 (0.0011)	0.6056 (0.0003)	0.6069 (0.0002)
1	0.5661 (0.0011)	0.6059 (0.0003)	0.6066 (0.0002)
2	0.6301 (0.0008)	0.7058 (0.0001)	0.7072 (0.0001)

4. Acknowledgement

I would like to thank Professor Thomas S. Ferguson for his many helpful suggestions and several discussions on this subject.

REFERENCES

Anderson, J. and Louis, T. (1979). The Properties of an Empirical Bayes Estimator Based on Dirichlet Process Hyper-priors. Preprint. Biostatistics Department, Harvard University.

Antoniak, C. E. (1974). Mixtures of Dirichlet Processes with Applications to Bayesian Nonparametric Problems. Ann. Statist., 2, 1152-1174.

Berger, J. (1980a). Improving on Inadmissible Estimators in Continuous Exponential Fimilies with Applications to Simultaneous Estimation of Gamma Scale Parameters. Ann. Statist., 8, 545-571.

Berger, J. (1980b). Statistical Decision Theory. Springer Verlag.

Berry, D. and Christensen, R. (1979). Empirical Bayes Estimation of a Binomial Parameter via Mixtures of Dirichlet Processes. Ann. Statist., 7, 558-568.

Ferguson, T. S. (1973). A Bayesian Analysis of Some Nonparametric Problems. Ann. Statist., 1, 209-230.

Ferguson, T. S. (1983). Bayesian Density Estimation by Mixtures of Normal Distributions, In: Rizvi, Rustagi and Siegmund, eds. Recent Advances in Statistics (Academic Press), 287-302.

Ghorai, J. K. and Susarla, V. (1982). Empirical Bayes Estimation of Probability Density Function with Dirichlet Process Prior, In: W. Grossman et al., eds., Probability and Statistical Inference (D. Reidel Co.), 101-114.

Kuo, L. (1985). Computations of Mixtures of Dirichlet Processes. SIAM Journal on Scientific and Statistical Computing. To appear.

Lo, A. Y. (1978). Bayesian Nonparametric Density Methods. Technical Report, University of California, Berkeley.

Lo, A. Y. (1984). On a Class of Bayesian Nonparametric Estimates: 1. Density Estimates. Ann. Statist., 12, 351-357.

Martz, H. and Waller, R. (1982). Bayesian Reliability Analysis. Wiley, New York.

Robbins, H. (1955). An Empirical Bayes Approach to Statistics. Proc. 3rd Berkeley Symp. Math. Statist. Prob., 157-163.

Robbins, H. (1980). An Empirical Bayes Estimation Problem. Proc. Natl. Acad. Sci., U.S.A., 77, 6988-6989.

Susarla, V. (1982). Empirical Bayes Theory. Encyclopedia of Statistical Sciences V.2., eds. Kotz, S., Johnson, N. and Read, C., Wiley, N.Y., 490-503.

END
DTIC
8-86